Breakpoint Characterization and Assessment for Position Effects in Patients with Tourette Syndrome and Rearrangements of Chromosome 18q22

Adam Charles Cuker

YALE UNIVERSITY

2003

Breakpoint Characterization and Assessment for Position Effects in Patients with Tourette Syndrome and Rearrangements of Chromosome 18q22

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

by

Adam Charles Cuker

2003

Abstract:

BREAKPOINT CHARACTERIZATION AND ASSESSMENT FOR POSITION EFFECTS IN TWO PATIENTS WITH TOURETTE SYNDROME AND REARRANGEMENTS OF CHROMOSOME 18Q22

Adam Cuker, Matthew W. State, John Greally, and David C. Ward. Department of Genetics, Yale University School of Medicine, New Haven, CT.

The objective of this work has been to identify a gene or genes involved in the pathogenesis of Tourette syndrome (TS), a neurodevelopmental disorder characterized by chronic motor and vocal tics. Multiple lines of evidence suggest that TS and a spectrum of related disorders including chronic tics and obsessive compulsive disorder (OCD) are genetically mediated, but research has yet to uncover a gene involved in disease etiology. The success of this research, consisting mainly of population genetic approaches, has likely been limited by a combination of genetic heterogeneity, uncertainty regarding genetic model parameters, and diagnostic ambiguity.

Molecular characterization of chromosomal rearrangements in rare affected patients with cytogenetic abnormalities represents an alternative strategy for disease gene identification. Two independently ascertained patients with TS spectrum phenotypes and rearrangements of chromosome 18q22 were identified and their cytogenetic abnormalities finely mapped. One of the rearrangements mapped to within approximately 800 kb of a previously reported 18q22 breakpoint in a TS spectrum pedigree. Assessment of the genomic interval defined by these breakpoints identified two nearby transcripts, neither of which were physically disrupted. Mutation screening

of these transcripts in 96 cytogentically normal TS spectrum patients did not reveal any nonsense or missense mutations.

Fluorescence *in situ* hybridization (FISH) replication timing studies were used to assess the epigenetic characteristics of the interval. These analyses showed significant replication asynchrony in one of the patients compared to controls. Replication was found to be relatively delayed across at least a 500 kb interval on the patient's abnormal chromosome compared to its normal homologue. These findings reflect broad epigenetic dysregulation in the region, and suggest that decreased or silenced expression of one or more genes in the interval may be contributing to the patient's phenotype.

Acknowledgements:

I am beholden to an exceptional group of people who provided me with invaluable assistance during the course of my research. First and foremost, I am grateful to Matthew State, who allowed me to take a leadership role in his project and provided me with the instruction and tools to carry out this role. I am truly fortunate to have worked with a physician-scientist so skilled in teaching, research, and patient care. He has been exactly the mentor I hoped to find when I entered medical school. I am equally grateful to David Ward, my thesis advisor, who generously gave of his time and tremendous wisdom during the project. I am privileged to have begun my career in his laboratory. I would also be remiss not to mention the contributions of John Greally, whose intellectual insights and practical guidance kept the project moving forward in the face of occasional adversity.

My colleagues in the Ward lab, Octavian Henigariu, Xiangyang Lu, Robert Reynolds, Xiaboa Zhong, Benjamin Sawicki, Jonas Lai, Alexandra Samms, Gretchen Graff, Manjunath Nimmakayalu, and Abby Donovan provided technical assistance and expertise through every step of the project. Even days of failed experiments were enjoyable thanks to their friendship and good humor.

I was fortunate to receive a tremendous amount of support for my research. Financial support came from medical student research training fellowships from the National Institutes of Health and the Howard Hughes Medical Institute. Statistical support came from Thomas Morgan, whose expertise was invaluable for the completion of this work. I must also acknowledge the contributions of Robert King of the Child

Study Center, whose clinical curiosity and acumen led to the identification of our patients.

To the patients and families that participated in this project, I am most indebted. If any advances in the field of Tourette syndrome are made as a result of this work, it is only because of their courage.

Finally, I am grateful to my own family. But for my parents' unwavering and unconditional support, I would never have been able to follow my dream of becoming a physician-scientist. And if not for my wife Amy's love, support, and zest for living, I doubt that I would have found this project or anything else in life nearly so inspiring.

Table of Contents:

Figures

Introduction:

Clinical features of TS

Tourette syndrome (TS) is a potentially debilitating neuropsychiatric disorder characterized by the presence of chronic motor and vocal tics. The disease usually has its onset in early childhood and follows a waxing and waning course. Motor tics are stereotyped fragments of normal motor movements. They may be simple, e.g. an eye blink or facial grimace, or more complex such as bending down and touching the floor. In severe cases, complex motor tics may take the form of copropraxia (obscene gesturing), echopraxia (mimicking another's movements), or self-injurious behaviors. Similarly, vocal tics may be classified as simple (e.g. throat clearing) or complex. In severe cases, complex vocal tics may manifest as coprolalia, echolalia, or palilalia (repetition of parts of words) (1, 2).

TS has a genetic basis

Evidence overwhelmingly suggests a genetic basis for TS (3). First-degree relatives of identified probands have roughly a 100-fold greater risk of TS than individuals in the general population (4, 5). Furthermore, twin studies have found concordance rates for TS to be greater than 50% for monozygotic twins and less than 10% for dizygotic twins (6, 7). When study methodology has involved direct patient examination and included the diagnosis of chronic tics as well as TS, monozygotic concordance has been determined to be as high as 100% (8).

Nevertheless, two decades of linkage analyses have not succeeded in identifying specific genes involved in disease pathogenesis (9). The most promising study to date from traditional linkage analysis demonstrated a LOD (logarithm of the odds) score of 3.24 on chromosome 11q23 (10). This finding awaits independent replication and identification of a disease-related gene.

Multiple factors including genetic heterogeneity, uncertainties regarding genetic model parameters, and diagnostic ambiguity have all likely hindered parametric linkage analyses of TS (11). For example, a particularly high degree of bilineality has been found in TS pedigrees (12, 13). This phenomenon involves mating between two affected but unrelated individuals. Pedigrees are sought in linkage studies in order to provide a relatively homogeneous study population. However, bilineality tends to produce families in which affected members have similar but genetically heterogeneous phenotypes.

Accurate assumptions regarding genetic parameters are also critical to the success of traditional linkage analyses. Most TS linkage studies have assumed an autosomal dominant mode of transmission with partial penetrance (14, 15), while more recent studies have challenged this assumption. Walkup and colleagues found that TS was most likely the result of a gene of major effect conferring more than half the overall risk for the disorder, with the remainder accounted for by genetic background and environmental factors (16). A recent segregation analysis of 108 extended families was unable to identify a Mendelian pattern of inheritance (11).

In addition, there is significant uncertainty regarding the population prevalence of TS. Estimates vary widely from a low of 5:1,000,000 to upwards of 1:100 among school age males (17). The figure most often cited places the prevalence at 5-10:10,000 (18).

Whether a TS gene mutation is assumed to be rare or relatively common has a significant impact on the outcome of linkage analysis.

Linkage studies are also very sensitive to determinations about what constitutes an affected individual. With regard to TS, this is a matter of considerable controversy. For instance, there is consensus that a significant degree of comorbidity exists between TS, Attention Deficit Hyperactivity Disorder (ADHD), and depression. However, there is no consensus on whether these disorders represent alternative patterns of expression due to a common genetic abnormality, as some investigators have suggested (19, 20). In addition, as discussed below, both obsessive compulsive disorder (OCD) and chronic tics in some but not all cases appear to constitute an alternative TS phenotype (21). Currently, this diagnostic ambiguity does not allow for differentiation between pedigree members with these findings who should and should not be considered affected in linkage studies.

As recognition of the obstacles facing traditional linkage studies has increased, non-parametric approaches to the study of TS have been undertaken. A sib-pair analysis by the Tourette Syndrome Association International Consortium for Genetics demonstrated two regions of the genome, one on chromosome 4q and another on 8p, that achieved LOD scores of greater than 2 (22). A recent association study of the Afrikaner population showed some evidence of association for a number of markers, though none corresponded to those identified in the Consortium study (23).

Evidence for a genetic relationship between TS, chronic tics, and OCD

While there is significant diagnostic uncertainty surrounding TS, there is general agreement that chronic tics and OCD, in some families, represent an alternative phenotypic expression of a common TS genotype. The rate of chronic tics and OCD in patients with TS is quite high. As many as 30 to 80 percent of TS probands meet diagnostic criteria for OCD, depending on how subjects are ascertained. These rates are considerably higher than the 2 to 3 percent prevalence of OCD in the general population (14). Moreover, there is a significantly increased risk for chronic tics and OCD in family members of TS patients. Probands presenting with TS alone have relatives with significantly elevated rates of chronic tics and/or OCD compared with matched controls (24, 16).

Rationale for the mapping of cytogenetic rearrangements

An alternative strategy for disease gene identification that is considerably less dependent on assumptions regarding genetic model parameters than traditional linkage analysis is the fine mapping of rare cytogenetic abnormalities in patients with TS and related phenotypes. This strategy has been successfully employed in mapping genes underlying a variety of single-gene disorders including mental retardation syndromes, neurological disorders, cardiovascular disease, and cancer syndromes (25). More recently, this approach has been used successfully in the investigation of neuropsychiatric and developmental disorders. In 1997, Kishino *et al.* identified a causative gene for

Angelman syndrome, in part through the identification of a patient with an inversion of chromosome 15 (26).

This strategy has also been applied to the study of TS in the hopes of identifying a gene or genes of major effect disrupted by a chromosomal rearrangement. A review of all published cases of cytogenetic abnormalities identified in TS patients reveals that three regions of the genome, on chromosomes 18q, 7q, and 8q, have been reported to be rearranged in more than one unrelated individual (27-30). Nonetheless, only one report to date has identified a structurally disrupted transcript (30), and its relevance to TS has yet to be established.

An important note regarding the fine mapping of cytogenetic abnormalities in TS patients is that it is unlikely to elucidate a common pathway in disease pathogenesis. Nevertheless, identification of even rare causal genetic abnormalities that are present in only a small subset of patients can be invaluable in elucidating physiologic pathways and genes that more commonly contribute to trait development.

18q abnormalities in two patients identified at Yale

Two independently ascertained patients with TS-spectrum phenotypes were identified in the Yale TS/OCD clinic, both of whom were found to have chromosomal abnormalities involving cytogenetic band 18q22. Case 1 involved a young man with OCD, chronic tics, and a paracentric inversion inv(18)(q21;q22). Apart from chronic motor tics and obsessive-compulsive symptoms including intrusive thoughts and an intense preoccupation with exactness, the patient was well developed, in good health, and had an above average IQ. The patient had been adopted from a country outside the

United States in infancy and his biological family was unknown. A family history was therefore not obtainable.

Case 2 involved a 14-year-old girl with a t(2;18)(p25;q22) translocation and severe OCD. She was otherwise healthy and exhibited normal development. The translocation was also found in the patient's father, who carried no psychiatric diagnosis and refused to undergo additional psychiatric evaluation. No other immediate family members, which included the patient's mother and a brother, carried the translocation or had any history of TS, OCD, or chronic tics.

Evidence for a gene or genes of etiologic importance on chromosome 18q

These patients' 18q22 breakpoints were of considerable interest, not only because of their proximity to each other, but also because of their proximity, as determined by conventional cytogenetic analysis, to a previously reported 18q22 breakpoint in a TS spectrum pedigree. In 1986, a TS patient with a t(7;18) translocation was reported (27). Analysis of first-degree relatives demonstrated that all of the proband's 4 siblings carried the translocation and that each showed evidence of tics and/or obsessive compulsive symptoms. In addition, the proband's father and two paternal aunts carried the translocation, though only one of the aunts was determined to be affected based on strict diagnostic criteria. Two cytogenetically normal paternal uncles were not affected. By conventional cytogenetic analysis, the chromosome 18 breakpoint in this pedigree was assigned to band 18q22.1. A decade later, this breakpoint was narrowed to a single YAC clone by somatic-cell hybrid and fluorescence *in situ* hybridization (FISH) analysis (28).

In addition to these three cases of 18q22 rearrangements, approximately 100 cases of 18q- syndrome have been reported in the literature. This rare disorder involves multiple congenital anomalies and, frequently, mental retardation, in association with large deletions of chromosome 18q distal to 18q21.1 (31, 32). Although tics are not a cardinal manifestation of the 18q- syndrome, case series have demonstrated a high rate of psychiatric symptomatology including aggression, impulsiveness, and attention difficulties (33). One patient with 18q- syndrome has been reported to have OCD (34). There have also been reports of movement disorders associated with the syndrome (35, 36). At present, it is unclear which gene or genes underlie the neuropsychiatric manifestations of the 18q- syndrome, though the *Myelin Basic Protein* gene on 18q23 has been suggested as a candidate. What is clear is that haplo-insufficiency of the distal 18q region is related to central nervous system disease, and that phenotypic expression of this disease is markedly heterogeneous (37).

Cytogenetic abnormalities, chromatin characteristics, and gene expression

There are several mechanisms by which a chromosomal rearrangement can disrupt a gene. The direct physical disruption of a gene's coding sequence is the most widely recognized of these. However, cytogenetic abnormalities may also have effects on genes distant from the site of rearrangement. Several cases have been described in which patients have clearly causative breakpoints as far as hundreds of kilobases away from the gene implicated in the etiology of their disease (38). There are several well-described mechanisms by which a chromosomal rearrangement may disrupt the function of distant genes. One possibility is that a rearrangement may disrupt a control element

lying hundreds to thousands of base pairs outside the coding sequence of a gene. A second possibility is that a rearrangement may lead to changes in the chromatin characteristics of a given genetic interval, thereby exerting position effects on genes within the interval and altering their expression.

The relationship between chromatin conformation and gene expression has been studied extensively in humans. Chromosomes are organized into structurally and functionally distinct domains known as chromatin domains. Transcriptionally active chromatin domains have the structural characteristics of euchromatin, i.e. they are relatively hypomethylated and hyperacetylated and exist in a loose uncondensed form. Conversely, non-expressing domains are characteristically heterochromatic, existing in a condensed form with a relatively high degree of methylation and low degree of acetylation.

A tight relationship has also been observed between the replication timing of a locus and its transcriptional activity (39). Expressed relatively euchromatic loci replicate early in S-phase whereas silent heterochromatic loci tend to replicate late. For example, housekeeping genes, which are active in all cells, replicate early. Tissue-specific genes, in contrast, tend to replicate early in those cells in which they are expressed and late in tissues in which they are transcriptionally silent (40).

The replication timing of loci characterized by mono-allelic expression has also been investigated. Studies carried out on imprinted loci have demonstrated replication asynchrony with the expressed allele replicating early and its silent homologue replicating late (41). Similarly, X-linked loci subject to X-chromosome inactivation have

been shown to replicate asynchronously, with the expressed allele replicating early in S-phase and its inactivated homologue replicating late (42).

While genomic imprinting and X-inactivation are normal epigenetic regulatory processes that demonstrate replication asynchrony, there are also examples in the literature describing chromosomal rearrangements that lead to significant changes in chromatin conformation, with a concomitant effect on gene expression and replication timing. The best characterized of these involve X;autosome translocations, in which the normal silencing mechanism of the X chromosome spreads to contiguous autosomal regions, transforming these regions from early replicating euchromatic DNA to late replicating heterochromatic DNA (43).

Several reliable assays exist by which to study replication timing as a marker for chromatin conformation and gene expression. The simplest and most widely used of these methods involves FISH hybridization of interphase nuclei. This technique allows for ready differentiation between replicated and unreplicated loci. Replicated loci are demonstrated by two closely approximated fluorescent signals (i.e., a doublet). Loci that have yet to undergo replication are represented by only one fluorescent signal (i.e., a singlet). The vast majority of homologous loci in the genome replicate synchronously, as evidenced by a large preponderance of singlet/singlet or doublet/doublet signals. However, those loci at which gene expression has been silenced on one allele replicate in an asynchronous fashion, which is manifested by an increased proportion of singlet/doublet signals. FISH analysis of interphase nuclei may thus be used to assay for replication asynchrony at a given locus, the presence of which suggests that one allele is expressing while its homologue is not.

While FISH hybridization of interphase nuclei is a simple and reliable assay for replication asynchrony, caution must be exercised in interpreting its results. Although it has been shown repeatedly that mono-allelic expression of a locus corresponds to replication asynchrony, the reverse does not always seem to be the case. For instance, in X;autosome translocations where the replication timing of the translocated autosome has clearly been affected, the expression of some but not all genes in the region has been found to be altered (43). A second caution is that chromosomal abnormalities as well as certain other diseases may alter replication timing not only in contiguous regions of a chromosome, but widely across the genome. Several studies have suggested that trisomy of a single chromosome may lead to genome-wide dysregulation of replication timing, and malignancy has likewise been shown to affect replication timing broadly (44-46).

Despite these caveats and limitations, FISH analysis of interphase nuclei is a confirmed strategy for identifying alterations in chromatin characteristics and gene expression. It allows for efficient evaluation of large genomic intervals prior to the identification of all transcripts in the region of interest. Furthermore, through the use of overlapping probes, it allows for determination of the extent and boundaries of replication asynchrony around a chromosomal breakpoint, thereby delineating a region around the breakpoint in which a disease gene or genes is most likely to reside.

Statement of purpose and hypothesis:

The overall objective of this work is to identify a gene or genes involved in the pathogenesis of TS and a spectrum of related phenotypes, including chronic tics and OCD, through the molecular characterization of rare cytogenetic abnormalities in patients with these phenotypes. This approach is based on the hypothesis that all of the identified abnormalities of chromosome 18q22 have resulted in structural or functional disruption of a common gene or genes and that this disruption, in turn, has produced similar neuropsychiatric phenotypes.

Methods:

The research described in this thesis is part of a larger ongoing project that began prior to my arrival in the laboratory and represents the collaborative efforts of several laboratory members. Elements of this project performed by others are noted as such in the text, and are included in order to place my work in a coherent and intelligible context.

Patients

Case 1. The patient was a 12 year-old boy of South Korean descent who was adopted by an American couple at the age of 3 months. He presented to the TS/OCD clinic at Yale for consultation after approximately 2 years of psychiatric treatment for obsessive compulsive symptoms, tics, depression, and anorexia. No data were available regarding prenatal or family history. The patient had a history of 2 febrile seizures as a toddler, each lasting 20-40 seconds and accompanied by fevers of 105-106°F. An EEG done several years later was unremarkable.

The patient had a normal developmental history. He was not noted to be dysmorphic on physical examination and demonstrated an above average full scale IQ (91^{st} percentile). Chromosomal testing revealed an inversion of chromosome inv(18)(q21;q22). Fragile X testing was negative. Routine laboratories including CBC, electrolytes, and thyroid function studies done at the time of his initial evaluation were unremarkable.

During his clinical evaluation at Yale, the patient was administered a Yale Global Tic Severity Scale (YGTSS), a clinician-rated semi-structured interview developed by

Leckman and colleagues for the purpose of evaluating the severity of tic symptoms (47). In addition, a self-report version of the YGTSS was completed by the patient's parents prior to his first visit.

In response to the YGTSS, the patient and his adoptive parents agreed that he developed eye-blinking at age 6, shoulder jerking at age 7, and teeth grinding and unwanted mouth movements beginning at age 8. Though the tics were described as mild and not leading to significant social morbidity, they were noted to be present on a daily basis over several years. The patient denied premonitory urges, but did describe a sense of relief after having tics. He also reported some feelings of physical tension in his joints prior to having tics, and endorsed "just right" phenomena. There was no history of vocal tics. The patient's parents noted a worst ever severity score of 6 out of 12, describing the motor tics as extremely forceful but relatively infrequent.

The patient's obsessive compulsive symptoms were evaluated using the Y-BOCS (Yale-Brown Obsessive Compulsive Scale). This standard diagnostic instrument includes a symptom checklist of commonly endorsed obsessions and compulsions, as well as a severity rating scale that assesses both current and worst-ever symptoms.

On the symptom checklist, the patient noted a fear of separation from a close family member, a fear of doing something embarrassing, hoarding and saving compulsions, an excessive preoccupation with right and wrong, a compelling need to know and to "ask and tell," fears of not saying the right thing, fears of losing things, excessive concern about his appearance, intrusive images, repeated checking for mistakes, mental rituals, and hair pulling.

The patient described his current obsessions as occupying more than 8 hours of his day. He noted that they were "extremely incapacitating," causing him "severe distress," and felt that he had little control over his obsessive thinking. With regard to his compulsions, he felt that they occupied between 3 and 8 hours of his day and were severely incapacitating. His total Y-BOCS score for current obsessions and compulsions was 30 out of a maximum of 40 points. At their worst, the obsessions were noted to occupy the patient's thinking for more than 8 hours a day and to cause extreme interference and distress. The patient did not feel that he had any control over this thinking. With respect to worst ever compulsions: these were performed for more than 8 hours a day, led to severe interference with daily functioning, and resulted in severe anxiety if the patient were prevented from carrying them out. His worst ever severity score was 34 out of a possible 40.

Based on the clinical presentation and scores on standardized instruments, the patient met criteria for moderately severe OCD and chronic motor tics.

Case 2. The patient was a 14 year-old girl evaluated for obsessive compulsive symptoms after a brief psychiatric hospitalization for a depressive episode. The patient's complaints were assessed in part using the clinician-administered CY-BOCS (Yale-Brown Obsessive Compulsive Scale, Child Version) that, like the adult version, consists of a symptom checklist as well as severity ratings for obsessions and compulsions. On the symptom checklist, the patient endorsed a pervasive preoccupation with and scrupulosity regarding schoolwork; intrusive thoughts unrelated to school that interrupted her; intrusive sounds, words or music; and a need for symmetry and exactness. She noted

repeatedly checking schoolwork for mistakes, rereading, erasing and rewriting, lining up objects or arranging them in a specific way, and hording school-related items.

The patient described her current obsessions as "extreme," occupying more than 8 hours of the day. She noted that they interfered "severely" with her social, academic, and family functioning, and that they caused "severe distress." She felt that she had no control over her obsessive thinking and that she "completely and willingly" yielded to all her obsessions. With regard to her compulsions, she felt that they too occupied more than 8 hours per day, caused severe to extreme interference with daily activities, and led to severe distress. As with her obsessions, the patient noted that she completely and willingly yielded to all of her compulsions and that she had little control over them. Her total Y-BOCS score for current obsessions and compulsions was 35 out of a maximum of 40 points. Based on her clinical history using the standardized instruments, the patient met diagnostic criteria for moderately severe OCD.

The patient had no evidence of tics by history or examination. She had suffered a single major depressive episode that required psychiatric hospitalization, and was diagnosed at discharge with major depression, OCD, and generalized anxiety disorder. There was no evidence of or history for psychotic thinking. The patient's medical history was remarkable for asthma and Lyme disease. Her developmental history was unremarkable and there was no evidence of dysmorphology on physical examination. Chromosomal testing demonstrated a translocation of chromosomes t(2;18)(p25q22). The patient's father was also noted to posses the translocation but did not carry a psychiatric diagnosis and refused psychiatric evaluation. No other family members had a history of OCD or tics.

Molecular mapping of chromosome 18q breakpoints

The chromosome 18 breakpoints in both patients were finely mapped using FISH as detailed below. Mapping of the telomeric inversion breakpoint in case 1 was completed prior to my arrival in the laboratory and is described elsewhere (48). I mapped the centromeric inversion breakpoint in case 1 as well as the translocation breakpoint in case 2 using the techniques described below.

Preparation of YAC DNA. 386 well plates containing CEPH B YACs were thawed at room temperature and 1 µl of YAC culture was taken from the appropriate well and placed in 50 ml of selective media in a culture flask. Cultures were grown at 30°C for 4-5 days with shaking (225 rpm). Once the cultures were opaque, they were transferred to 50 ml conical Falcon tubes and centrifuged at 6,000 x g for 10 minutes. After decanting, the pellet was resuspended in 1 ml of SCEM buffer (75 µl 2-mercaptoethanol and 925 µl of SCE buffer consisting of 0.9M sorbitol, 0.1M sodium citrate, 0.06M EDTA pH 8.0). The solution was transferred to a 1.5 ml Eppendorf tube, covered with paraffin, and rocked at 37°C for 2 hours. The tube was then centrifuged for 5 minutes at 10,000 x g in a cold room and the supernatant gently poured off. 333 µl of 50mM Tris/20mM EDTA was added and the pellet was resuspended. 33 µl of 10% SLS was added and mixed well by inversion. The Eppendorf tube was again sealed with paraffin and placed in a 60°C water bath for 20 minutes. 133 µl of 5M KOAc (pH 4.5) was added and the tubes were incubated on ice for 60 minutes. The tubes were centrifuged at 10,000 x g at room temperature for 10 minutes. The supernatant containing DNA was then poured into a fresh Eppendorf tube and 1 ml of 100% ethanol was added. The precipitate was centrifuged at 10,000 x g at 4°C. After decanting the

supernatant, the pellet was washed in 1 ml of 70% ethanol and then resuspended in 50 µl of TE (pH 8.0).

Preparation of BAC DNA. 2 ml of LB media with 25 µl/ml chloramphenicol were seeded with BACs taken from a RPCI-11 library stored at -80°C. A 20 µl disposable pipette tip was gently streaked across the frozen culture aliquot and then deposited into 15 ml Falcon tubes containing the culture medium. These were grown overnight with agitation at 37°C. 5 ml of this solution was subsequently streaked on LB/agarose/chloramphenicol plates prepared using standard methods (49) and grown overnight. Single colonies were selected and used to seed 50 ml of LB with chloramphenicol and grown for 18 hours. The 50 ml cultures were centrifuged at 6,000 x g for 20 minutes. After pouring off the supernatant, the remaining pellet was resuspended in 10 ml of solution S1 containing 50mM Tris-HCl, 10mM EDTA, 100 µg/ml RNase A (Nucleobond DNA Purification Kit, Clontech: Palo Alto, CA). After fully resuspending the pellet, 10 ml of S2 solution (200mM NaOH and 1% SDS) was added and the tubes were gently inverted 5 times. The suspension was incubated at room temperature for 2 minutes. 10 ml of ice-cold S3 solution (2.8M KOAc, pH 5.1) was subsequently added and again the tubes were gently inverted 5 times and incubated on ice for 5 minutes. The lysate was then poured into funnels containing manufacturer-supplied filters that had been pre-wetted with 1 ml of equilibration buffer N2 (100mM Tris, 15% ethanol, 900mM KCl, pH 6.3). The solution was filtered through a Nucleobond AX100 column that had also been pre-equilibrated with 5 ml of N2 buffer. The flow-through was discarded and the DNA-containing column was washed thrice, each time with 4 ml of fresh wash solution N3 (100mM Tris, 15% ethanol, 1.15M KCl, pH 6.3). The column

was then eluted with 5 ml of elution buffer N5 (100mM Tris, 15% ethanol, 1M KCl, pH 8.5) preheated to 50°C. 900 µl of solution containing purified BAC DNA was aliquoted into 1.5 ml Eppendorf tubes and precipitated with 600 µl of isopropanol. The tubes were then spun at 10,000 x g in a cold room. After centrifugation, the supernatant was poured off and the remaining pellet was washed with 1 ml of 70% ethanol and resuspended in 10 µl of TE (pH 8.0).

Probe labeling using nick translation. For each probe, a solution was prepared containing 1 µg of YAC or BAC DNA mixed with 5 µl of 10X nick translation buffer, 5 µl of 0.5mM dATP/dGTP/dCTP, 5 µl of 0.5mM haptene-conjugated dUTP (biotin-dUTP or digoxigenin-dUTP), 5 µl of 0.1M BME, 3 µl of 10U/µl DNase diluted 1:1000, and 2 µl of 3U/µl DNA polymerase I. The solution was then diluted with water to a total volume of 50 µl and incubated at 15°C for 90 minutes. The quality of the nick translation was tested by removing 5 µl and boiling the aliquot for 5 minutes in a screw-top 1.5 ml tube. The sample was then placed on ice for 3 minutes and loaded with loading buffer and ethidium bromide on a 2% agarose gel. The presence of DNA fragments from 200-400 base pairs in length was considered acceptable. If the obtained fragments were too long, additional DNase and DNA polymerase I were added and the reaction was continued at 15°C. An aliquot was checked again after 15-30 minutes depending on the status of the original nick. Once an adequate nick translation was achieved, the remainder of the reaction was stopped with 2 µl of 0.5M EDTA and 1 µl of SDS heated at 68°C. The labeled probes were then stored at -70°C until use.

Probe precipitation with COT1 DNA. In a 1.5 ml Eppendorf tube, 100 ng of nicked BAC or YAC DNA was added to 1 volume of 1 µg/µl salmon sperm DNA, 1

volume of 1 μg/μl COT1 DNA, 0.1 volumes of 3M NaOAc, and 2.5 volumes of 100% ethanol. The mixture was incubated either overnight at 4°C or for 1 hour at -80°C. The solution was then centrifuged at 10,000 x g in a cold room for 15 minutes. The supernatant was decanted and the pellet dried. The pellet was then resuspended in 5 μl of 80% formamide/15% dextran sulfate. In some cases, 2 probes with different conjugated haptenes were co-precipitated using the same reaction conditions.

Slide preparation. Metaphase chromosomes were prepared from human peripheral blood lymphocytes stimulated with phytohemagglutinin and harvested 72 hours later from unsynchronized cultures by standard procedures. Cells were spread on glass slides and chemically aged per the method described by Henigariu and colleagues (50).

Hybridization and washing. The probe suspension was shaken vigorously for 20 minutes. After shaking, the probe(s) was placed in a 75°C heat block for 10 minutes and then allowed to pre-anneal in a 37°C heat block for 1 hour. Slides were denatured simultaneously with 150 μl of 70% formamide/2X SSC and heating to 75°C for 90 seconds followed by immediate cooling to 16°C using a slide block on a PCR cycler. The denatured slides were placed in ice cold 70% ethanol for 5 minutes and then transferred to ice cold 90% and 100% ethanol baths for 5 minutes each, after which they were set aside to air dry. The probe mixture was added to a pre-marked region of the slide containing adequate metaphase spreads. A 20 mm x 20 mm cover slip was applied and secured in an airtight fashion using rubber cement around the edges of the cover slip. The slides were then placed in a humidifying chamber overnight at 37°C.

The following day, the slides were removed from the moist chamber. While holding the cover slip in place, the rubber cement was removed with forceps. With the cover slip still in place, the slide was submerged in a Coplin jar containing 30 ml of Wash 1 (50% low grade formamide/2X SSC) at 42°C for 5 minutes with gentle agitation to allow the cover slip to fall off. The slides were then washed 3 times for 5 minutes each with fresh Wash 1 at 42°C. They were washed a fourth time with Wash 2 (0.1X SSC) at 60°C for 5 minutes.

Slides were tapped free of excess solution and 200 µl of blocking solution (3% BSA/4X SSC/0.1% Tween-20) was placed in the region of the probe. The slides were covered with 24 mm x 60 mm cover slips and placed back in the humidifying chamber for 30 minutes at 37°C. During this time, avidin-FITC (1:400 dilution) and anti-digoxigenin-rhodamine (1:150 dilution) were added to detection buffer (1%BSA/4X SSC/0.1% Tween-20) to make detection solution. Cover slips were again removed from the slides and 200 µl of detection solution was added to each slide. New 24 mm x 60 mm cover slips were placed on the slides and bubbles were removed by gentle tapping with forceps. The covered slides were placed again in the humidifying chamber at 37°C for 30 minutes. Upon removal from the moist chamber, slides were subjected to 3 5-minute washes in Wash 3 (0.1% Tween-20/4X SSC) at 42°C with gentle shaking so that the cover slips would fall off. The slides were then counterstained with DAPI (4,6-diamidino-2-phenylindoledihydrochloride, 200 ng/ml) for 4 minutes and then submerged in a Coplin jar containing 2X SSC for 2 minutes. 1 drop of Vectashield antifade was added and the cover slip replaced. Slides were then ready for visualization. A CCD camera (PM512, Photometrics) was used to visualize fluorescent signals. Grayscale

images were obtained sequentially for fluorescein, rhodamine, and DAPI with precision filter sets (Zeiss) to minimize the image shifts. The grayscale images were pseudocolored and merged.

Identification of putative coding sequence

Identification of putative coding sequence in the vicinity of the 18q22 breakpoints was carried out by another member of the laboratory using a variety of approaches including human-human BLAST homology searches, mouse-human sequence comparison, EST database searches, and gene-prediction algorithms. Expression of putative coding regions was assessed by PCR amplification from various tissue-specific cDNA libraries and reverse transcriptase PCR using several tissues as template. A comprehensive account of the methodology employed is detailed elsewhere (48).

Heteroduplex analysis

Mutation screening of identified putative coding sequence was undertaken by another member of the laboratory using heteroduplex analysis (48).

Replication timing studies

FISH was carried out using BAC-derived probes as described above with the following modifications: (1) Cell cultures were pulsed with 10μM bromodeoxyuridine (BrdU) for 90 minutes prior to harvesting in order to allow for incorporation into newly synthesizing DNA. (2) Three color fluorescence imaging was performed. One BAC probe was labeled using dUTP-11-digoxigenin and detected with anti-digoxigenin-

rhodamine (1:150 dilution). A second BAC probe was direct-labeled with dUTP-DEAC. The BrdU-pulsed S-phase cells were detected with a 1:10 dilution of anti-BrdU-FITC (Pharmingen). (3) Interphase nuclei were counted with respect to the pattern of hybridization present at each locus. As noted, three patterns of hybridization were observed, a singlet/singlet (SS) pattern, a doublet/doublet (DD) pattern, and a singlet/doublet (SD) pattern. Only those interphase nuclei showing BrdU incorporation as a marker for S-phase were counted. (4) Two co-localizing BAC probes were used in each experiment to distinguish signals that were representative of true hybridization as well to differentiate the abnormal chromosome from its normal homologue. (5) Each slide was counted independently by two individuals in the laboratory. Whenever possible, the identity of the probes and the source of the cell material were not known to the rater. Certain experiments, for instance those involving breakpoint-spanning probes, precluded blinding. Inter-rater reliability was calculated using a simple Kappa test. (6) Several slides containing cells from a single control were hybridized at different times with the identical probe to ensure reproducibility of hybridization and counting. (7) Statistical analysis of SD percentages from the replication timing experiments was undertaken using a chi-square test with 1 degree of freedom. All reported significance levels are two-tailed.

Results:

Molecular mapping of breakpoints

Identification of spanning YAC clones. Mapping of the breakpoints was undertaken prior to the availability of relevant draft sequence from the human genome project. Databases at the National Center for Biotechnology Information were used to identify sequence tagged sites (STSs) mapping to the 18q21 and 18q22-23 regions (http://www.ncbi.nlm.nih.gov/genome/sts/). The website of the Whitehead Institute (www.genome.wi.mit.edu) was subsequently used to identify YAC clones corresponding to the STSs and mapping to the region of interest. A single contig (WC 18.4) was noted to contain YAC clones spanning the chromosome 18q21-23 region. Representative YACs from across this contig were selected and obtained from a CEPH B YAC library. These clones were used in FISH hybridizations with lymphocytes from each patient.

The mapping strategy involved localization of YAC clones to a locus either proximal or distal to the breakpoints, thereby allowing for progressive narrowing of the genomic intervals of interest until spanning YACs were identified for each breakpoint. Expected FISH results employing this strategy are shown for case 1 (Fig. 1) and case 2 (Fig. 2). In both cases, hybridization of multiple YACs led to the identification of clones spanning the chromosome 18q breakpoints. In case 1, YAC 846A2 was found to span the 18q22 breakpoint, and YAC clone 804B10 was found to span the 18q21 breakpoint. For case 2, the YAC spanning the 18q22 translocation breakpoint was found to be clone 766F12 (data not shown).

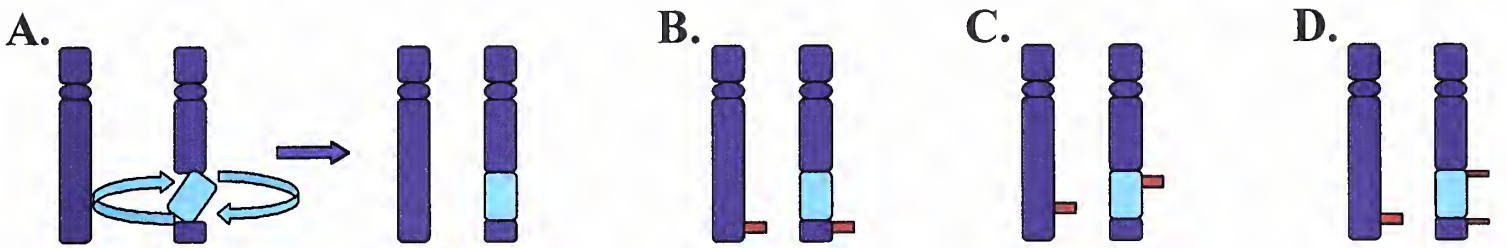

Fig. 1A. Expected FISH results for mapping the telomeric aspect of a paracentric inversion. (A) A paracentric inversion involving the long arm of chromosome 18 is shown. The light blue portion represents the inverted segment of the abnormal chromosome. (B) Expected FISH results from a probe (red rectangle) that maps telomeric to the inversion breakpoint are depicted. The hybridization signal on the normal and inverted chromosomes is in the same location. (C) Expected results from a probe that maps just centromeric to the breakpoint are shown. The signal on the inverted chromosome is visualized at a point proximal to where it is seen on the normal chromosome. (D) Expected FISH results from a probe spanning the inversion breakpoint are demonstrated. Probe hybridizing to the abnormal chromosome splits to yield signals at both the centromeric and telomeric aspects of the inversion, while the normal chromosome yields a signal only at the telomeric aspect.

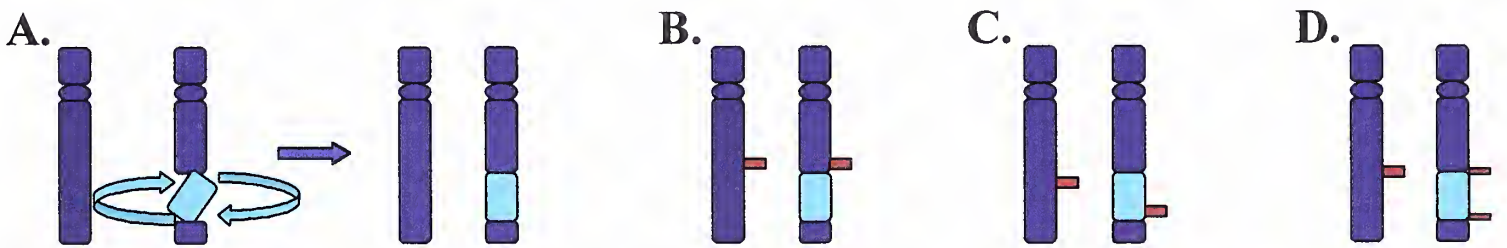

Fig 1B. Expected FISH results for mapping the centromeric aspect of a paracentric inversion. (A) A paracentric inversion involving the long arm of chromosome 18 is shown. The light blue portion represents the inverted segment of the abnormal chromosome. (B) Expected FISH results from a probe (red rectangle) that maps centromeric to the inversion breakpoint are demonstrated. The position of the probe on the normal and inverted chromosomes is the same. (C) Expected results from a probe that maps just telomeric to the breakpoint are depicted. The hybridization signal on the abnormal chromosome is shifted distal to its location on the normal chromosome. (D) Expected FISH results from a probe spanning the inversion breakpoint are shown. Probe hybridizing to the abnormal chromosome splits to yield signals at both the centromeric and telomeric aspects of the inversion, while the normal chromosome yields a signal only at the centromeric aspect.

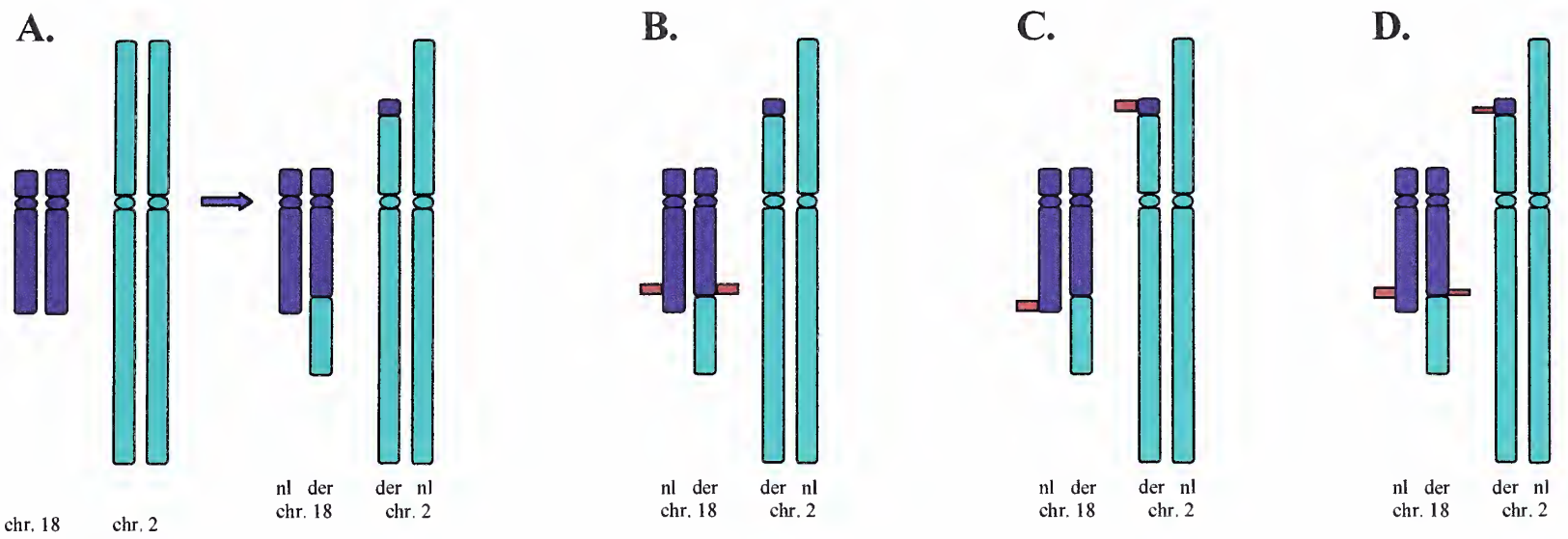

Fig. 2. Expected FISH results for mapping the chromosome 18 aspect of a t(2;18) (p25;q22) translocation.
(A) Normal chromosomes 18 and 2 are shown in blue and green, respectively. Also shown is a t(2;18)(p25;q22) translocation. Chromosome 2 material translocated to the der(18) is shown in green. Chromosome 18 material translocated to the der(2) is shown in blue. (B) Expected results for a probe (red rectangle) that maps centromeric to the chromosome 18 breakpoint are depicted. Hybridization signal is seen at the same position on both the normal and derivative chromosomes 18. (C) Expected results for a probe that maps telomeric to the chromosome 18 breakpoint are demonstrated. Hybridizations are seen on the normal chromosome 18 and on the der(2). (D) Expected FISH results from a probe that spans the chromosome 18 breakpoint are shown. Signals appear on the normal and der(18) as well as on the der(2).

Identification of spanning BAC clones. Once spanning YACs were found, BAC clones corresponding to these spanning YACs were identified. For case 1, several approaches including the screening of a RPCI-11 BAC library with radio-labeled EST and STS markers, clone walking using TAIL (thermal asymmetric interlaced) PCR, and further library screening with radio-labeled non-repetitive end sequences were employed in the search for spanning BAC clones. This search was performed by another member of the lab and is detailed elsewhere (48).

As additional draft sequence was made available through sequencing of the human genome, alternative methods for identifying BAC clones mapping to the region of interest became available. The High Throughput Genome Sequencing (HGTS) and Genome Sequence Survey (GSS) databases, in combination with the BLAST search engine (http;//www.ncbi.nlm.nih.gov/blast/htmlblastcgihelp.html#nucleotide_databases), provided the means for virtual clone-walking. STS and EST markers of interest were first analyzed to identify and mask any repetitive sequence using RepeatMasker (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker). Unique sequence was then used for nucleotide BLAST searches (http://www.ncbi.nlm.nih.gov/blast). As draft sequences from BAC clones were identified, they were screened for repeat sequences and unique regions were once again used for BLAST searches. This method resulted in the construction of several putative BAC contigs in the regions surrounding the 18q22 breakpoints in cases 1 and 2. The positions of these BACs were subsequently confirmed with FISH and spanning BACs were identified using the same strategy as was employed in the identification of spanning YAC clones. The overlapping BAC clones 600G22 and 644A7 were found to span the 18q22 inversion breakpoint in case 1 (Fig. 3A). For case

2, BAC clones 240N5 and 88B2 were determined to span the 18q22 translocation breakpoint (Fig. 3B).

Delineation of a 5.5 Mb interval on chromosome 18q22 by three cytogenetic breakpoints. The position of the translocation breakpoint identified by Boghosian-Sell and colleagues (28) was approximated based on the finding that marker WI-3559 mapped telomeric to this breakpoint (J. Overhauser, unpublished data). The sequence information made available as a result of the human genome project (http:genome.ucsc.edu and http://www.ncbi.nih.nlm.gov) allowed for an approximation of the intervals between the t(7;18)(q22-31;q22) translocation identified by Boghosian-Sell *et al.* and the two rearrangements identified at Yale. As shown in Fig. 4, the smallest distance between any two of these three breakpoints, approximately 800 kb, separated the telomeric inversion breakpoint in case 1 and the previously characterized t(7;18) translocation breakpoint (28).

Identification of putative coding sequence

The genomic interval containing the 18q22 breakpoints was extensively analyzed for coding sequence by another member of the lab using BLAST EST homology searches, comparison with mouse sequence, and various gene prediction algorithms (48).

A single EST was identified on BAC 600G22 mapping approximately 100 kb telomeric to the inversion breakpoint. The clone was subsequently identified in a fetal thymus library and named *GTSCR-1* (*Gilles de la Tourette Syndrome Chromosomal Region 1*) (Fig. 4). Upon sequencing, it was found to contain 3 exons and encode a

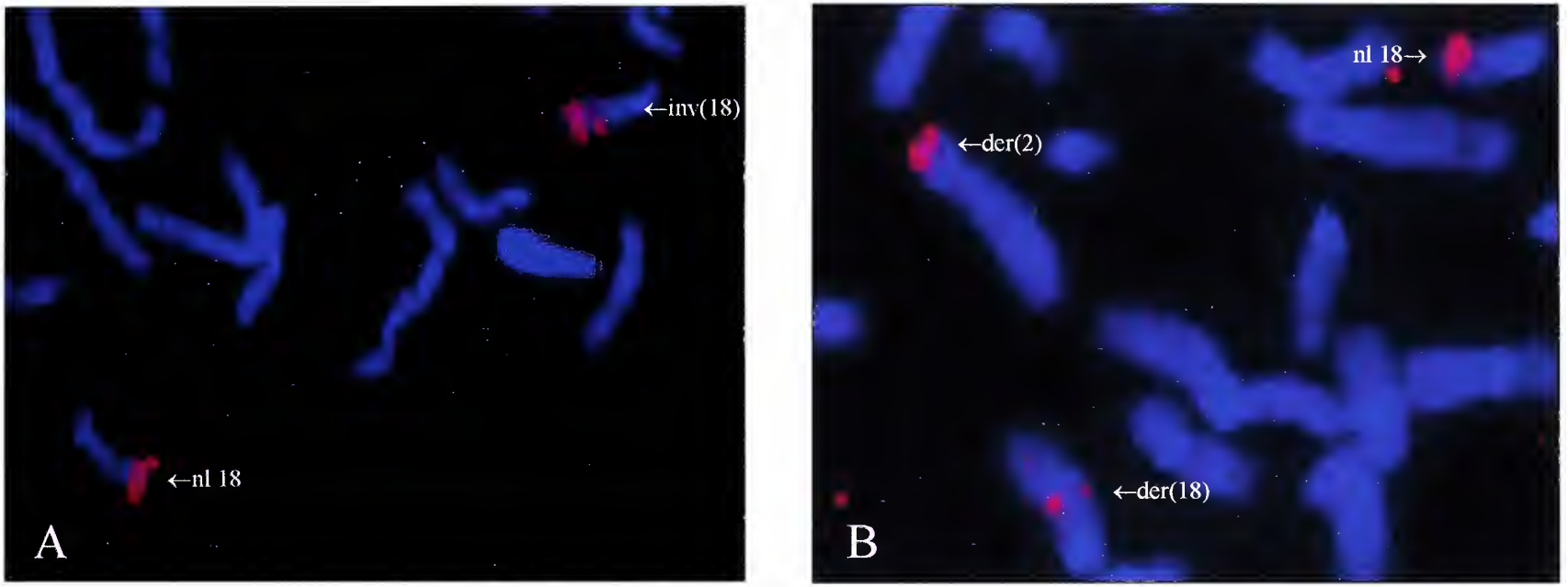

Fig. 3. FISH mapping of two rearrangements of chromosome 18q22. (A) A probe derived from BAC 600G22 spans the telomeric breakpoint in a patient with a paracentric inversion of chromosome 18q (case 1). The probe hybridizes to two regions on the inv(18) and only one region on the normal chromosome (nl 18). (B) A probe derived from BAC 240N5 spans the chromosome 18 breakpoint in a patient with a t(2;18) translocation (case 2). The probe hybridizes to the der(2) and der(18) chromosomes as well as the normal chromosome 18 (nl 18).

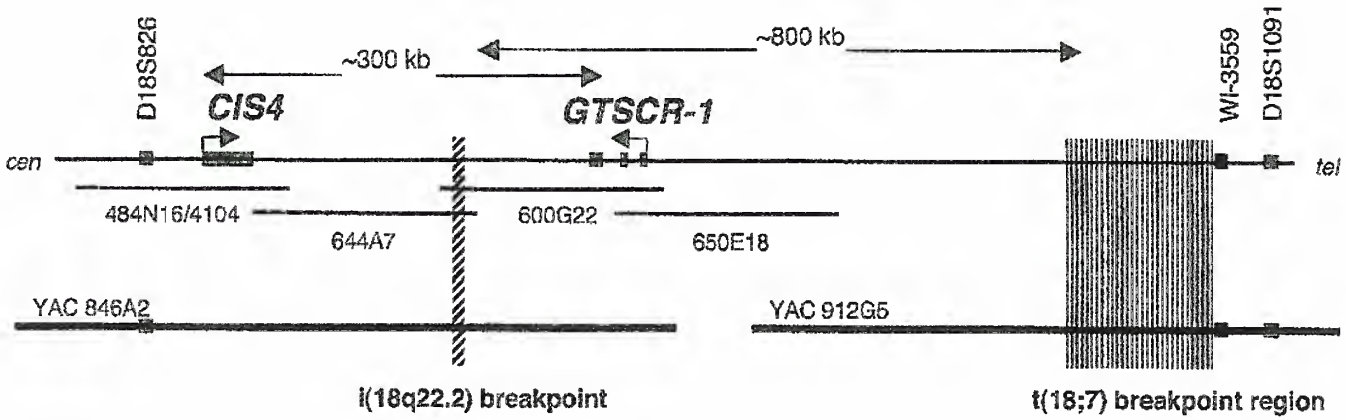

Fig. 4. Results of FISH mapping of both YAC and BAC clones in the region around the 18q22 inversion breakpoint. BACs are represented by thin horizontal lines and YACs are represented by heavy black lines at the bottom of the diagram. The approximate position of the inv(18)(q21;q22) inversion breakpoint is noted in the region of overlap between BAC clones 644A7 and 600G22. The approximate position of the previously noted t(18;7)(q22;q22-31) translocation breakpoint (28) is marked by the gray box on the right of the diagram. The t(18;2)(q22;p25) translocation breakpoint in case 2 was localized to a position approximately 4.7 Mb centromeric to the inversion breakpoint and is not shown on the diagram. Two transcripts were identified in the vicinity of the 18q22 breakpoint. *CIS4* was mapped to BAC 41O4 and *GTSCR-1* was mapped to BAC clones 600G22 and 650E18. Approximate distances separating the two transcripts and the two breakpoints are noted.

putative product of 138 amino acids. It showed no significant homology to other proteins, ESTs, or protein motifs.

A second putative coding region was found on BAC 41O4, mapping approximately 150 kb centromeric to the inversion breakpoint. Translation of the longest open reading frame in this region was found to be identical to the known gene *CIS4* (*cytokine inducible SH-2 containing protein 4*) (Fig. 4).

Mutation screening

dHPLC evaluation of coding regions and intron-exon splice junctions for the transcripts *CIS4* and *GTSCR-1* was performed by another member of the lab. An analysis of 96 cytogenetically normal patients with TS-spectrum phenotypes revealed no missense or nonsense mutations (48).

Assessment of replication timing at the 18q22 breakpoint in case 1

FISH analyses of interphase nuclei were used to assess for the presence and degree of replication asynchrony at various loci. Loci that replicate synchronously show either a single hybridization on each homologue (singlet/singlet or SS pattern) or two hybridization signals on each homologue (doublet/doublet or DD pattern), depending on whether the locus is pre- or post-replication. Loci at which one homologue has replicated and the other has not are visualized as a singlet/doublet (SD) pattern of hybridization (Fig. 5).

To assess the degree of replication asynchrony around the 18q22 breakpoint in case 1, the spanning BAC 600G22 was hybridized to interphase nuclei from the patient

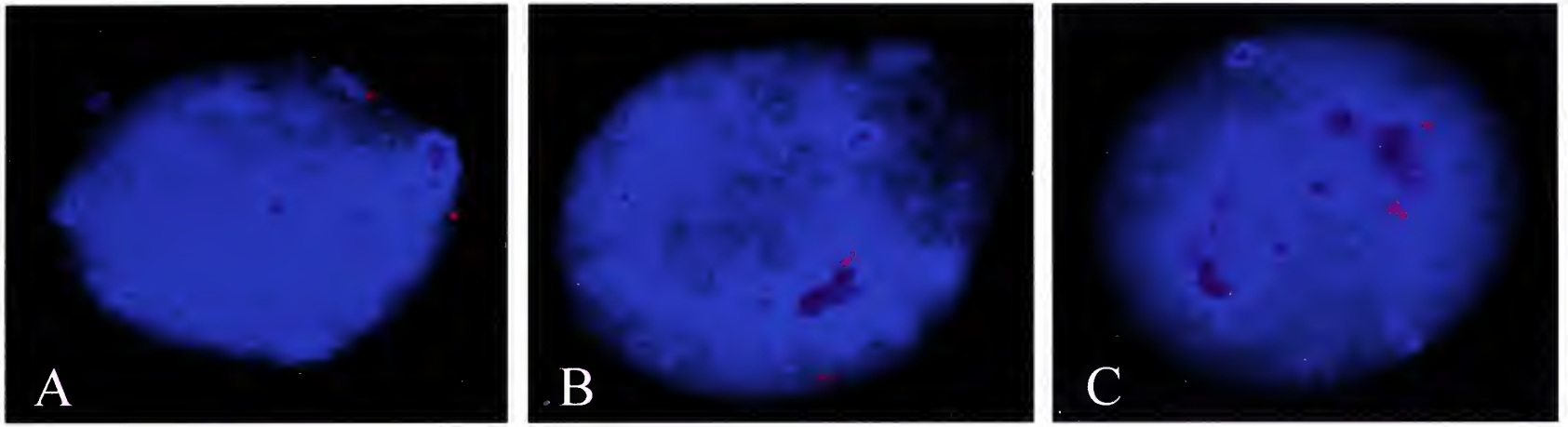

Fig. 5. Representative hybridization patterns obtained in replication timing studies using FISH probes on interphase nuclei. Panels A, B, and C demonstrate the singlet/singlet (SS) pattern, doublet/doublet (DD) pattern, and singlet/doublet (SD) pattern visualized on interphase nuclei, respectively.

and the nuclei were counted with respect to their pattern of hybridization. In order to assess baseline replication timing in the region, 600G22 was also hybridized to S-phase nuclei from an unaffected, cytogenetically normal individual. A BAC from the SNRPN-SNURF locus, an imprinted locus mapping to the Prader-Willi region of chromosome 15, was hybridized to patient cells from case 1 as a control for replication asynchrony. A BAC from the synchronously replicating CFTR region was likewise hybridized to patient cells as a control for replication synchrony. For each hybridization, a second nearby BAC probe was co-hybridized to confirm hybridization of the experimental probe. Observation was limited to nuclei in S-phase, as distinguished by BrdU pulsation and detection with anti-BrdU-FITC (Fig. 6). For each slide, at least 60 adequately hybridized S-phase nuclei were counted.

Statistical analysis was carried out using a chi square test with 1 degree of freedom. The percentage of SD hybridizations found in the patient for BAC 600G22 was significantly higher than that for the same probe in a cytogenetically normal cell line (0.45 vs. 0.27, $p<0.01$) (Fig. 7). Similarly, the SD percentage at the breakpoint was significantly higher than that for a probe mapping to the CFTR control locus in the same patient (0.45 vs. 0.24, $p<0.005$) (Fig. 7). In contrast, the rate of asynchrony at the breakpoint was not different from that at the asynchronously replicating Prader-Willi locus in the same patient (0.45 vs. 0.44, $p>0.5$) (Fig. 7).

A broader region around the inversion breakpoint was assessed with additional BAC probes 644A7 and 650E18, which overlap clone 600G22 on its centromeric and telomeric flanks, respectively (Fig. 4). These clones were hybridized to

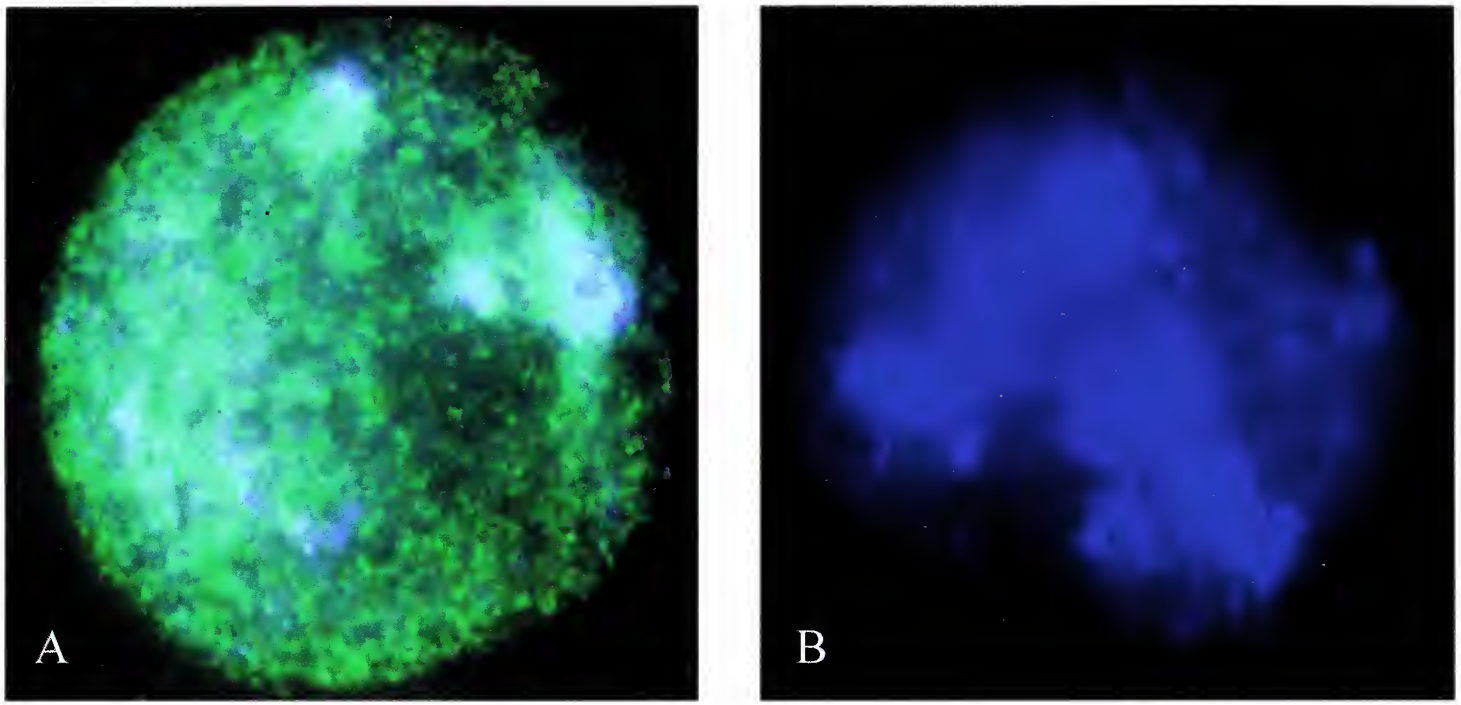

Fig. 6. Identification of S-phase nuclei with BrdU pulsation. Cells were pulsed with BrdU prior to harvesting and subsequently detected with anti-BrdU FITC. (A) An interphase nucleus exhibits a characteristic stippled pattern upon detection with anti-BrdU FITC, signifying that it is in S-phase. (B) An interphase nucleus not in S-phase shows no evidence of BrdU incorporation.

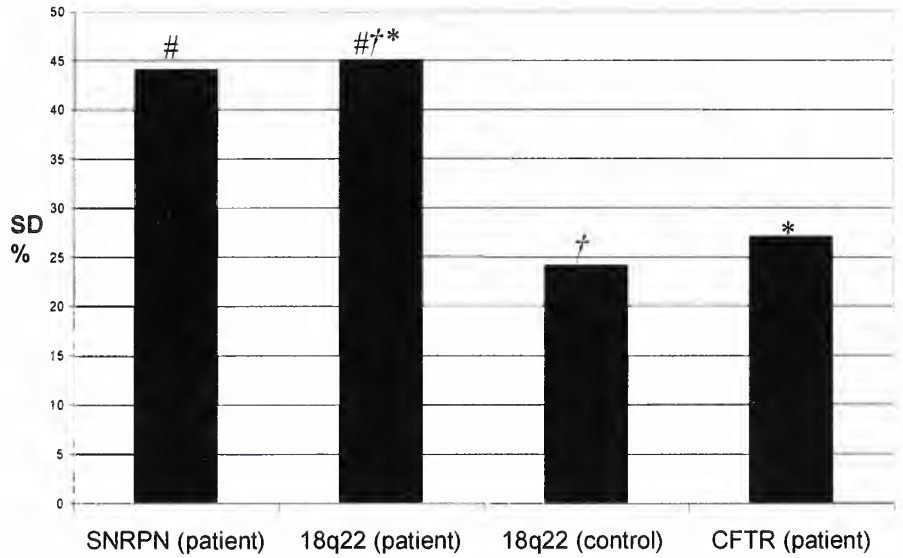

Fig. 7. Percentage ofsinglet/doublet signals for 4 conditions. Lane 1: A BAC probe containing the gene SNURF-SNRPN known to replicate asynchronously is hybridized to patient cells. Lane 2: BAC probe 600G22 spanning the patient's inversion breakpoint is hybridized to patient cells. Lane 3: BAC 600G22 is hybridized to cytogenetically normal control cells. Lane 4: A BAC probe mapping to the CFTR gene known to replicate synchronously is hybridized to patient cells. For each condition, 100 interphase nuclei were counted. Inter-rater reliability was calculated using the kappa statistic and showed excellent agreement (k=0.9583). A chi-square test with 1 degree of freedom was carried out comparing the 18q probe hybridized to the patient cell line versus control cells (lanes 2 vs 3). The difference was significant (†p<0.01). A comparison of SD percentages at the 18q locus in the patient and at the CFTR locus in the patient (lanes 2 vs 4) also showed a significant difference (*p<0.005). The SD percentages at the SNRPN locus and at the 18q22 locus in the patient (lanes 1 vs. 2) were indistinguishable (#p>0.5).

interphase nuclei from the patient as well as to a cytogentically normal control cell line. In each instance, the degree of replication asynchrony in the patient was found to be significantly higher than that for the control cells. For BAC 644A7, the proportion of SD hybridizations for patient and control cells was 0.40 and 0.21, respectively ($p<0.05$). For BAC 650E18, the SD proportion for patient compared to control nuclei was 0.52 versus 0.27 ($p<0.01$). These findings define at least a 500 kb region of abnormal replication, presumably due to the patient's chromosomal rearrangement (Fig. 8).

The patient's normal and abnormal chromosomes 18 were distinguishable in interphase nuclei via the use of FISH probes lying very close to, but on opposite sides of, the inversion breakpoint. With this probe selection, there was co-localization of hybridization signals on the normal chromosome, but marked separation of signals on the abnormal chromosome (Fig. 9). An analysis of the BAC probe spanning the inversion breakpoint demonstrated that 81% of cells found to replicate asynchronously in the patient had the singlet (representative of later replication) on the chromosome bearing the inversion compared to 19% on the normal homologue ($p<0.0001$) (Fig. 10).

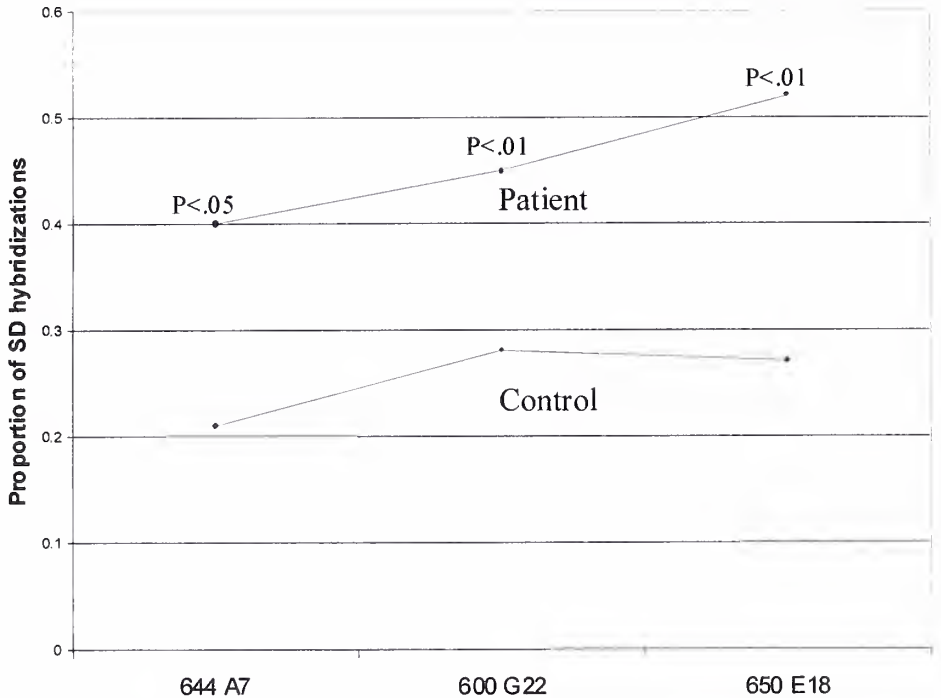

Fig. 8. Replication timing in the region surrounding the 18q22 breakpoint. The frequency of singlet/doublet signals for 3 BACs in the region surrounding the 18q22 inversion breakpoint are shown. The top line represents data from the patient cells. The bottom line shows data from controls. The total number of interphase nuclei counted for each slide were as follows: BAC 644A7, patient=60, control=60; BAC 600G22, patient=129, control 88; and BAC 650E18, patient=60, control=60. All differences between control and experimental conditions were significant and are noted in their respective positions on the graph.

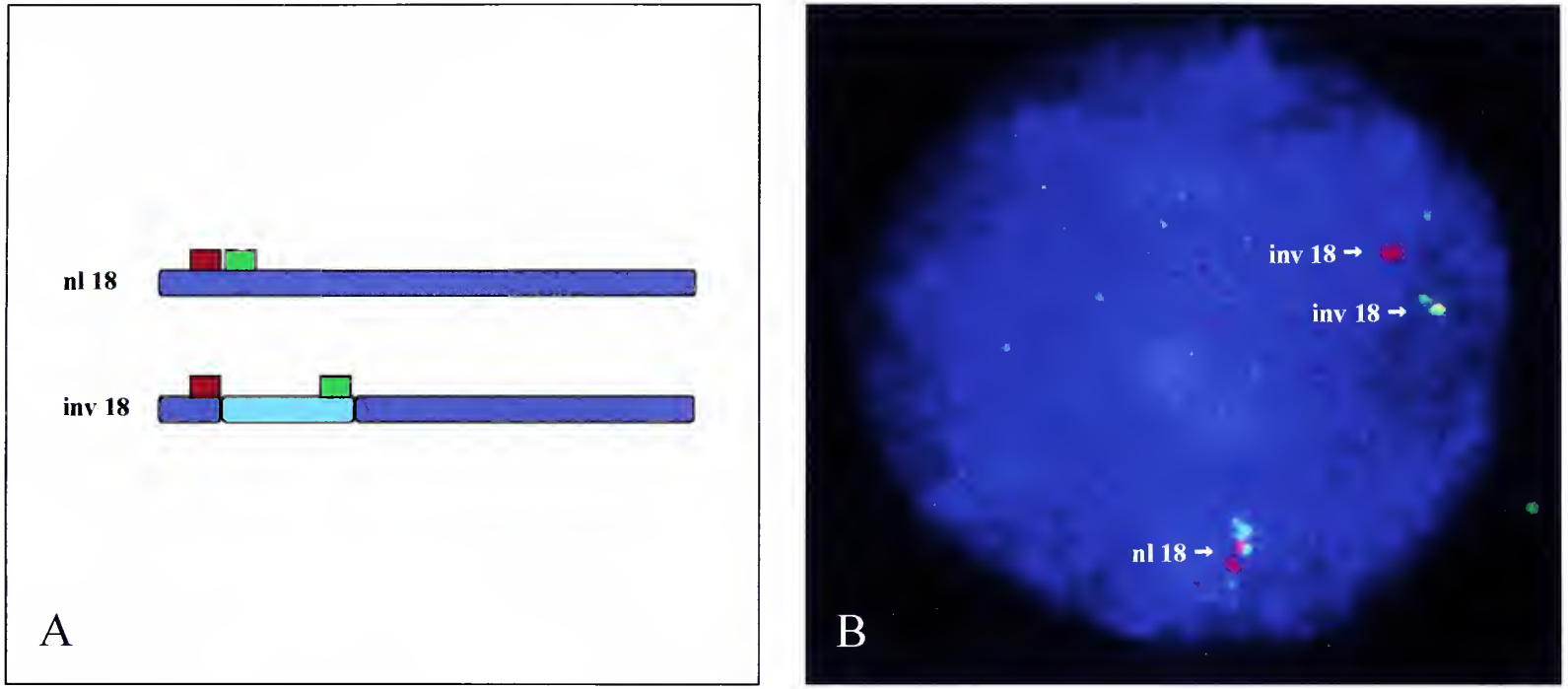

Fig. 9. Differentiating the normal from the inverted chromosome 18 in replication timing studies. (A) A paracentric inversion of chromosome 18 (case 1) is shown with the inverted segment in light blue. The patient's normal chromosome 18 (nl 18) and inverted chromosome 18 (inv 18) are distinguishable in interphase nuclei by the use of two FISH probes lying just telomeric (red rectangle) and just centromeric (green rectangle) to the distal inversion breakpoint. On the normal chromosome, these probes remain adjacent. On the inverted chromosome, the probes are split apart to yield spatially separated signals. (B) An interphase nucleus has been hybridized with probes derived from BACs 650E18 (red signal) and 484N16 (green signal), which lye just telomeric and just centromeric to the distal inversion breakpoint, respectively (Fig. 4). The signals co-localize on the normal chromosome (nl 18), but split apart on the inverted chromosome (inv 18), allowing for differentiation between the normal and abnormal chromosomes 18 in interphase nuclei.

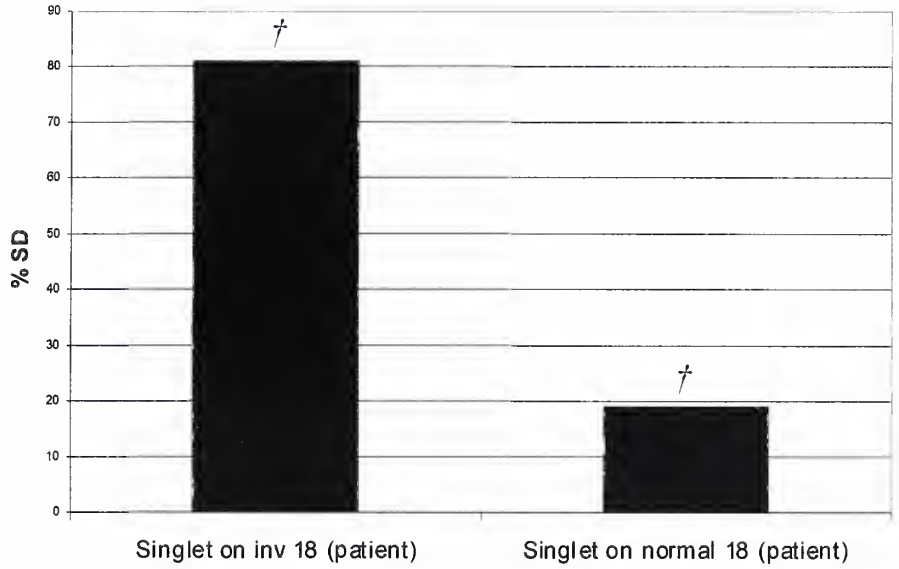

Fig. 10. Percentage of singlet signals on the normal versus inverted chromosome 18. BAC 600G22 was hybridized to patient cells. 81% of the singlet/doublet hybridizations observed (47/58) showed the singlet on the inverted chromosome (lane 1) while only 19% of singlets were found on the normal chromosome (lane 2). The difference was significant (†$p<0.0001$).

Discussion:

Two patients with rearrangements of chromosome 18q22 and TS spectrum phenotypes were identified. The fine mapping of these rearrangements using YAC and BAC-derived FISH probes delineated a 5.5 Mb region of chromosome 18q22 containing three TS-associated rearrangement breakpoints. The smallest distance between any two of these breakpoints, approximately 800 kb, separated the telomeric inversion breakpoint in case 1 and the previously characterized t(7;18)(q22-31;q22) translocation breakpoint (28). Analysis of the genomic interval did not reveal a structurally disrupted gene, but did identify two nearby transcripts, *CIS4* and the newly identified *GTSCR-1*. The search for additional genes in the region will continue as data from the Human Genome Project becomes increasingly refined.

The known gene, *CIS4*, mapped approximately 150 kb centromeric to the case 1 distal inversion breakpoint (Fig. 4). Although the function of CIS4 is not well understood, it is believed to play a role in cytokine-induced signaling in the immune system. Consistent with this hypothesis, *CIS4* expression has been noted to be high in several cytokine-dependent hematopoetic cell lines including factor-independent chronic myelogenous leukemia (CMK) and human erythroleukemia (hel) (51).

A possible role for CIS4 in TS pathogenesis is appealing in light of existing hypotheses that postulate an immune-mediated mechanism for the disorder. Specifically, a post-infectious etiology following Group A beta-hemolytic streptococcal (GABHS) infection has been hypothesized. GABHS is known to trigger immune-mediated sequelae including acute rheumatic fever and post-streptococcal

glomerulonephritis in genetically susceptible individuals (52). One of the major clinical criteria for the diagnosis of rheumatic fever, Sydenham's chorea, presents with choreiform movements of the distal limbs. In a substantial number of cases, patients may also manifest frank motor or vocal tics as well as obsessive compulsive symptoms (53). In addition to similarities in their clinical presentations, TS, OCD, and Sydenham's chorea appear to share common anatomic targets in the central nervous system, specifically the basal ganglia and related circuits (54). Taken together, the parallels between these disorders suggest that, at least in some cases, they may share a common etiopathogenesis. Additional clinical and epidemiologic studies are needed to confirm the association between the development of TS spectrum phenotypes and antecedent GABHS infection.

Much less is known about the novel gene *GTSCR-1*. A range of library screens indicates that it is a rare transcript with expression confirmed only in human thymus (2). Neither sequence homologies to known genes nor similarities to any functional protein motifs were identified and, as such, the function of GTSCR-1 remains unknown. However, its expression in thymus raises the possibility that it too may subserve a role in the immune system.

As neither *CIS4* nor *GTSCR-1* were found to be structurally disrupted by the identified rearrangements, FISH replication timing analyses were performed to determine if the chromosomal inversion in case 1 might have altered the epigenetic properties of the region, potentially leading to altered expression of one or more genes in the vicinity. Marked replication asynchrony about the breakpoint and extending over an interval of at least 500 kb was identified, with evidence for a relative delay in replication on the

abnormal chromosome compared with its normal homologue. Similar analysis of a cytogenetically normal unaffected control demonstrated no evidence of asynchrony in the region. In addition, expected patterns of replication timing were upheld in patient cells at loci expected to demonstrate either asynchronous (SNRPN-SNURF) or synchronous (CFTR) replication.

The relative delay in replication observed on the patient's rearranged chromosome 18 is consistent with reduced transcriptional activity in the region, and suggests decreased or silenced expression of one or more nearby genes. Given the negative findings on heteroduplex analysis of 96 cytogenetically normal TS-spectrum patients (48), it is unlikely that mutations in either the *CIS4* or *GTSCR-1* genes represent a common cause of TS. However, the identification of even a rare TS gene or genes could provide important insights into the genetic and physiologic mechanisms that underlie the disorder. Further exploration of the functions of CIS4 and GTSCR-1 might assist in the elucidation of such mechanisms. Moreover, the putative role of one or both of these genes in the immune system and the hypothesized role of GABHS infection in TS pathogenesis suggests the importance of additional mutation screening in TS spectrum patients with a clear history of antecedent GABHS infection.

Our findings on replication timing analysis also suggest a number of avenues of future inquiry, many of which are already underway. First, BACs mapping to each of the known genes in the interval of interest will be used as FISH probes in future replication timing studies in an effort to gauge the extent to which these genes have been subject to position effects. Second, the BAC contig extending in both directions from the case 1 inversion breakpoint will continue to be assessed with respect to replication timing in

order to delineate the boundaries of asynchrony. Third, studies of replication timing on the soon-to-be available patient material from the previously identified t(7;18) translocation (28) will be undertaken. Similar analysis of replication timing at and around the chromosome 18q22 breakpoint in case 2 has already commenced. Mapping and comparing the full extent of replication asynchrony for all three cases may assist in defining a candidate interval in which to narrow the search for a gene or genes of etiologic significance. Fourth, replication timing studies will be carried out in cytogenetically normal patients with TS spectrum phenotypes in order to screen for epigenetic dysregulation in the region of interest in this population. Fifth, approaches for confirming the relevance of our replication timing to the epigenetic properties of the region of interest will be pursued including methylation studies and direct comparisons of patient versus control gene expression in any transcripts that can be quantified in accessible tissue.

Use of replication timing as an assay for epigenetic phenomena has important advantages in the study of chromosomal abnormalities. First, it allows for investigation of large regions of the genome compared with alternatives such as methylation-sensitive restriction enzymes or bisulfite sequencing. The ability to readily screen regions of interest in 100-200 kb segments is particularly valuable in the identification and mapping of long-range position effects. In addition, FISH replication timing studies may be carried out in peripheral lymphocytes, regardless of whether these cells express the genes under study. Although a gene's absolute replication timing within a given tissue is a function of its expression profile in that tissue, synchrony of replication occurs independent of cell type. Use of lymphocytes allows for the ready investigation of

epigenetic phenomena involving genes that may be expressed in inaccessible tissue, an asset that seems particularly valuable in the study of neuropsychiatric illness.

Despite these advantages and the confirmed use of FISH methodology for examining replication timing, several important issues regarding FISH replication timing analysis merit further discussion. First, there is little consideration of statistical issues with respect to this technique in the current literature (2). Investigators have, by and large, employed a "rule of thumb" approach to data collection and interpretation. For instance, one or a small number of slides per condition are generally evaluated and 100 nuclei per slide are counted. Among those nuclei, a rate of 30% SD signals is considered indicative of replication asynchrony. In an effort to determine the adequacy of this approach, the statistical significance of the results was assessed using a chi-square test. Power calculations suggested that approximately 60 interphase nuclei per condition would need to be counted in order to ascertain a 20% difference in the proportion of SD signals between synchronously and asynchronously replicating loci.

A second important methodological caution pertains to the percentage of SD hybridizations found at negative control loci. A baseline SD rate of approximately 10-15% has been noted in the literature for bi-allelically expressed loci in normal cells. This relatively small rate of SD hybridization is thought to be due to inefficient hybridization and/or imaging artifact rather than true replication asynchrony (55). Our methods, though seemingly identical to those published in the literature, routinely identified a 20-25% SD rate at bi-allelically expressing control loci. Explanations for this discrepancy are not obvious. One possibility is that our counting rules, while internally consistent as demonstrated by excellent inter-rater reliability, tended to undercount doublets in favor of

SS and SD hybridization patterns. This possibility is being further investigated by comparing all SS, SD, and DD counts at a number of loci. Nevertheless, the degree of difference we detected in SD percentages between synchronously replicating and asynchronously replicating control loci, approximately 15-20%, is completely consistent with that found in the literature.

In addition to the limitations associated with assays of replication timing, the clinical characterization of our patients and their relevance to the study of TS is worthy of discussion. The absence of clear biologic markers for TS and many other neuropsychiatric disorders demands that clinician-researchers rely heavily on history and presentation in establishing a diagnosis. In an effort to address this issue, standardized diagnostic instruments including the Y-BOCS and YGTSS, discussed above, have been developed.

Despite the marked improvement in diagnostic schemes brought about by the use of these standardized instruments, there remain instances in which diagnostic issues are particularly problematic for genetics research. With respect to mild or sub-clinical cases, for example, the boundary between normal and pathologic behavior can be difficult to define. The notion of functional impairment has become central to most diagnostic schemes in an effort to clarify this boundary. Attempts to measure the subjective distress and functional consequences of tics and obsessive compulsive phenomena have thus become central elements of standardized diagnostic approaches. The two patients identified at Yale were both found to have high levels of distress and functional impairment using the aforementioned standardized instruments. These findings suggest

that the clinical phenomena observed in these patients represent true psychopathology rather than normal variants.

Despite the limitations in our methodology, the clustering of three TS spectrum-associated, independently ascertained, chromosome 18q22 rearrangements suggests that the region is a promising candidate for containing a gene or genes of etiologic importance. Our finding in case 1 of long-range position effects in the 18q22 interval provides a model for how rearrangements mapping to non-coding regions of the genome and located hundreds of kb apart from one another may contribute to the production of similar phenotypes. Furthermore, our detection of regional epigenetic changes in a patient with a TS-related phenotype and a chromosomal rearrangement suggests a novel mechanism for neuropsychiatric pathogenesis that may be worth pursuing in related disorders. The search for additional patients with TS spectrum phenotypes and cytogenetic abnormalities continues in the hopes that molecular characterization of patient breakpoints may someday lead to the identification of a TS gene.

References:

1. Leckman, J.F., and Cohen, D.J. 1999. Evolving models of pathogenesis. In *Tourettes Syndrome: Tics, Obsessions and Compulsions*. J.F. Leckman and D.J. Cohen, editors. New York: John Wiley and Sons. 155-175.

2. State, M.W. 2001. *Evaluation of a Chromosome 18 Locus Associated with Tourette Syndrome Phenotypes: Breakpoint Characterization, Transcript Assessment, and Candidate Gene Analysis*. New Haven: Yale University Press. 155 pp.

3. State, M.W., Lombroso, P.J., Pauls, D.L., and Leckman, J.F. 2000. The genetics of childhood psychiatric disorders: a decade of progress. *J. Am. Acad. Child Adolesc. Psychiatry*. 39:946-962.

4. Pauls, D.L., Raymond, C.L., Stevenson, J.M., and Leckman, J.F. 1991. A family study of Gilles de la Tourette syndrome. *Am. J. Hum. Genet*. 48:154-163.

5. Pauls, D.L., Alsobrook, J.P., Gelertner, J., and Leckman, J.F. 1999. Genetic vulnerability. In *Tourette's Syndrome: Tics, Obsessions, Compulsions*. J.F. Leckman, editor. New York: John Wiley and Sons. 194-212.

6. Hyde, T.M., Aaronson, B.A., Randolph, C., Rickler, K.C., and Weinberger, D.R. 1992. Relationship of birth weight to the phenotypic expression of Gilles de la Tourette's syndrome in monozygotic twins. *Neurology*. 42:652-658.

7. Price, R.A., Kidd, K.K., Cohen, D.J., Pauls, D.L., and Leckman, J.F. 1985. A twin study of Tourette syndrome. *Arch. Gen. Psychiatry*. 42:815-820.

8. Walkup, J.T., Leckman, J.F., Price, R.A., Resnick, S., Hardin, M. *et al.* 1987. Non genetic factors associated with the expression of Tourette's syndrome. *Proceedings of the 34th Annual Meeting of the American Academy of Child and Adolescent Psychiatry.*

9. Barr, C.L., Wigg, K.G., Pakstis, A.J., Kurlan, R., Pauls, D., *et al.* 1999. Genome scan for linkage to Gilles de la Tourette syndrome. *Am. J. Med. Genet.* 88:437-445.

10. Merette C., Brassard, A., Potvin, A., Bouvier, H., Rousseau, F. *et al.* 2000. Significant linkage for Tourette syndrome in a large French Canadian family. *Am. J. Hum. Genet.* 67:1008-1013.

11. Seuchter, S.A., Hebebrand, J., Klug, B., Knapp, M., Lehmkuhl, G. *et al.* 2000. Complex segregation analysis of families ascertained through Gilles de la Tourette syndrome. *Genet. Epidemiol.* 18:33-47.

12. Kurlan, R., Eapen, V., Stern, J., McDermott, M.P., and Robertson, M.M. 1994. Bilineal transmission in Tourette's syndrome families. *Neurology.* 44:2336-2342.

13. McMahon, W.M., van de Wetering, B.J., Filloux, F., Betit K., Coon, H. *et al.* 1996. Bilineal transmission and phenotypic variation of Tourette's disorder in a large pedigree. *J. Am. Acad. Child Adolesc. Psychiatry.* 35:672-680.

14. Pauls, D.L. and Leckman, J.F. 1986. The inheritance of Gilles de la Tourette's syndrome and associated behaviors. Evidence for autosomal dominant transmission. *N. Engl. J. Med.* 315:993-997.

15. Price, R.A., Pauls, D.L., Kruger, S.D., and Caine, E.D. 1988. Family data support a dominant major gene for Tourette syndrome. *Psychiatry Res.* 24:251-261.

16. Walkup, J.T., LaBuda, M.C., Singer, H.S., Brown, J., Riddle, M.A. *et al.* 1996. Family study and segregation analysis of Tourette's syndrome: Evidence for a mixed model of inheritance. *Am. J. Hum. Genet.* 59:684-693.

17. Robertson, M.M., and Stern, J. 1998. Tic disorders: new developments in Tourette syndrome and related disorders. *Curr. Opin. Neurol.* 11:373-380.

18. Apter, A., Pauls, D.L., Bleich, A., Zohar, A.H., Kron, S. *et al.* 1993. An epidemiologic study of Gilles de la Tourette's syndrome in Israel. *Arch. Gen. Psychiatry.* 50:734-738.

19. Comings, D.E., and Comings, B.G. 1984. Tourette's syndrome and attention deficit disorder with hyperactivity: are they genetically related? *J. Am. Acad. Child Psychiatry.* 23:138-146.

20. Pauls, D.L., Hurst, C.R., Kruger, S.D., Leckman, J.F., Kidd, K.K. *et al.* 1986. Gilles de la Tourette's syndrome and attention deficit disorder with hyperactivity. Evidence against a genetic relationship. *Arch. Gen.Psychiatry.* 43:1177-1179.

21. Pauls, D.L., Towbin, K.E., Leckman, J.F., Zahner, G.E., and Cohen, D.J. 1986. Gilles de la Tourette syndrome and obsessive-compulsive disorder. Evidence supporting a genetic relationship. *Arch. Gen. Psychiatry.* 43:1180-1182.

22. Consortium IMGSoA. 1998. A full genome scan for autism with evidence for linkage to a region on chromosome 7q. *Hum. Mol. Genet.* 7:571-578.

23. Simonic, I., Gericke, G.S., Ott, J., and Weber, J.L. 1998. Identification of genetic markers associated with the Gilles de la Tourette syndrome in an Africaner population. *Am. J. Hum. Genet.* 63:839-846.

24. Pauls, D.L., Leckman, J.F., Towbin, K.E., Zahner, G.E., and Cohen, D.J. 1986. A possible genetic relationship exists between Tourette's syndrome and obsessive-compulsive disorder. *Psychopharmacol. Bull.* 22:730-733.

25. Collins, F.S. 1992. Postional cloning: let's not call it reverse any more. *Nat. Genet.* 1:3-6.

26. Kishino, T., Lalande, M., and Wagstaff, J. 1997. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat. Genet.* 15:70-73.

27. Comings, D.E., Comings, B.G., Dietz, G., Muhlman, D., Okada, T. *et al.* 1986. Evidence the Tourette syndrome gene is at 18q22.1. In *Human Genetics: Proceedings of the 7th International Congress of Human Genetics.* Vogel, F., and Sperling, K., editors. Berlin: Springer. 620.

28. Boghosian-Sell, L., Comings, D.E., and Overhauser, J. 1996. Tourette syndrome in a pedigree with a 7;18 translocation: identification of a YAC spanning the translocation breakpoint at 18q22.3. *Am. J. Hum. Genet.* 59:999-1005.

29. Matsumoto, N., David, D.E., Johnson, E.W., Konecki, D., Burmester, J.K. *et al.* 2000. Breakpoint sequences of an 1;8 translocation in a family with Gilles de la Tourette syndrome. *Eur. J. Hum. Genet.* 8:875-883.

30. Petek, E., Windpassinger, C., Vincent, J.B., Cheung, J., Boright, A.P. *et al.* 2001. Disruption of a novel gene (*IMMP2L*) by a breakpoint in 7q31 associated with Tourette syndrome. *Am. J. Hum. Genet.* 68:848-858.

31. de Grouchy, J., Royer, P., Ch, S., and M, L. 1964. Deletion partielle des bras longs du chromosome 18. *Pathol. Biol.* 12:579-582.

32. Wilson, M.G., Towner, J.W., Forsman, I., and Siris, E. 1979. Syndromes associated with deletion of the long arm of chromosome 18[del(18q)]. *Am. J. Med. Genet.* 3:155-174.

33. Mahr, R.N., Moberg, P.J., Overhauser, J., Strathdee, G., Kamholz, J., *et al.* 1996. Neuropsychiatry of the 18q- syndrome. *Am. J. Med. Genet.* 67:172-178.

34. Donnai. 1987. Gene location in Tourette syndrome. *Lancet.* 14:627.

35. Miller, G., Mowrey, P.N., Hopper, K.D., Frankel, C.A. and Ladda, R.L. 1990. Neurologic manifestations in 18q- syndrome. *Am. J. Med. Genet.* 37:128-132.

36. Gordon, M.F., Bressman, S., Brin, M.F., de Leon, D., Warburton, D. *et al.* 1995. Dystonia in a patient with deletion of 18q. *Mov. Disorders.* 10:496-499.

37. Strathdee, G., Sutherland, R., Jonsson, J.J., Sataloff, R., Kohonen-Corish, M. *et al.* 1997. Molecular characterization of patients with 18q23 deletions. *Am. J. Hum. Genet.* 60:860-868.

38. Kleinjan, D.J., and van Heyningen, V. 1998. Position effect in human genetic disease. *Hum. Mol. Genet.* 7:1611-1618.

39. Gunaratne, P.H., Nakao, M., Ledbetter, D.H., Sutcliffe, J.S., and Chinault, A.C. 1995. Tissue-specific and allele-specific replication timing control in the imprinted human Prader-Willi syndrome region. *Genes Dev.* 9:808-820.

40. Goldman, M.A., Holmquist, G.P., Gray, M.C., Caston, L.A., and Nag, A. 1984. Replication timing of genes and middle repetitive sequences. *Science*. 224:686-692.

41. Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I. *et al.* 1993. Allele-specific replication timing of imprinted gene regions. *Nature*. 364:459-463.

42. Boggs, B.A., and Chinault, A.C. 1994. Analysis of replication timing properties of human X-chromosomal loci by fluorescence *in situ* hybridization. *Proc. Natl. Acad. Sci. U S A.* 91:6083-6087.

43. White, W.M., Willard, H.F., Van Dyke, D.L., and Wolff, D.J. 1998. The spreading of X inactivation into autosomal material of an X;autosome translocation: evidence for a difference between autosomal and X-chromosomal DNA. *Am. J. Hum. Genet.* 63:20-28.

44. Amiel, A., Avivi, L., Gaber, E., and Fejgin, M.D. 1998. Asynchronous replication of allelic loci in Down syndrome. *Eur. J. Hum. Genet.* 6:359-364.

45. Amiel, A., Korenstein, A., Gaber, E., and Avivi, L. 1999. Asynchronous replication of alleles in genomes carrying an extra autosome. *Eur. J. Hum. Genet.* 7:223-230.

46. Amiel, A., Litmanovitch, T., Lishner, M., Mor, A., Gaber, E. *et al.* 1998. Temporal differences in replication timing of homologous loci in malignant cells derived from CML and lymphoma patients. *Genes Chromosom. Cancer*. 22:225-231.

47. Leckman, J.F., Riddle, M.A., Hardin, M.T., Ort, S.I., Swartz, K.L. *et al.* 1989. The Yale global tic severity scale: initial testing of a clinician-rated scale of tic severity. *J. Am. Acad. Child Adolesc. Psychiatry.* 28:566-573.

48. State, M.W., Greally, J., Cuker, A., Bowers, P., Henegariu, O. *et al.* 2003. Epigenetic abnormalities associated with a chromosome 18(q21-q22) inversion and a Gilles de la Tourette syndrome phenotype. *Proc. Natl. Acad. Sci. U S A.* In press.

49. Sambrook, J., and Russell, D.W. 2001. *Molecular cloning: a laboratory manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

50. Henegariu, O., Hereema, N.A., Wright, L., Bray-Ward, P., Ward, D.C. *et al.* 2001. Improvements in cytogenetic slide preparation: controlled chromosome spreading, chemical aging and gradual denaturing. *Cytometry.* 43:101-109.

51. Masuhara, M., Sakamoto, H., Matsumoto, A., Suzuki, R., Yasukawa, H. *et al.* 1997. Cloning and characterization of novel *CIS* family genes. *Biochem. Biophys. Res. Commun.* 239:439.

52. Bisno, A.L. 1991. Group A streptococcal infection and acute rheumatic fever. *N. Engl. J. Med.* 325:783-93.

53. Swedo, S.E., Rapoport, J.L., Cheslow, D.L., Leonard, H.L., Ayoub, E.M. *et al.* 1989. High prevalence of obsessive-compulsive symptoms in patients with Stdenham's chorea. *Am. J. Psychiatry.* 146:246-9.

54. Mercadante, M.T., Campos, M.C., Marques-Dias, M.J., Miguel, E.C. and Leckman, J. 1997. Vocal tics in Sydenham's chorea. *J. Am. Acad. Child Adolesc. Psychiatry.* 36:305-6.

55. Selig, S., Okumura, K., Ward, D. C. & Cedar, H. 1992. Delineation of DNA replication time zones by fluorescence *in situ* hybridization. *EMBO J.* 11:1217-25.

HARVEY CUSHING/JOHN HAY WHITNEY MEDICAL LIBRARY

MANUSCRIPT THESES

Unpublished theses submitted for the Master's and Doctor's degrees and deposited in the Medical Library are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but passages must not be copied without permission of the authors, and without proper credit being given in subsequent written or published work.

This thesis by
has been used by the following person, whose signatures attest their acceptance of the above restrictions.

NAME AND ADDRESS **DATE**